**Miguel-Chinchilla, L., Heasley, E., Loiselle, S. and Thornhill, I. (2019) 'Local and landscape influences on turbidity in urban streams: a global approach using citizen scientists', *Freshwater Science*, 38 (2), pp. 303-320.**

Official URL: https://doi.org/10.1086/703460

# ResearchSPAce

http://researchspace.bathspa.ac.uk/

This published version is made available in accordance with publisher policies. Please cite only the published version using the reference above.

Your access and use of this document is based on your acceptance of the ResearchSPAce Metadata and Data Policies, as well as applicable law:- https://researchspace.bathspa.ac.uk/policies.html

Unless you accept the terms of these Policies in full, you do not have permission to download this document.

This cover sheet may not be removed from the document.

Please scroll down to view the document.

# Local and landscape influences on turbidity in urban streams: a global approach using citizen scientists

**Leticia Miguel-Chinchilla[1,4], Eleanore Heasley[2,5], Steven Loiselle[1,6], and Ian Thornhill[1,3,7]**

[1]Earthwatch Institute (Europe), Mayfield House, 256 Banbury Road, Summertown, Oxford, OX2 7DE, UK
[2]Department of Geography, King's College London, Aldwych, London, WC2B 4BG, UK
[3]College of Liberal Arts (CoLA), Bath Spa University, Newton St Loe, Bath, BA2 9BN, UK

**Abstract:** The ecological degradation of urban rivers and streams has been termed the 'urban stream syndrome' and attributed to increased catchment urbanization. Limiting future degradation requires an understanding of the drivers of reduced water quality at both catchment and site scales. The goal of this study was to identify the probable drivers of turbidity in river ecosystems in highly urbanized areas, under the premise that turbidity does not respond consistently to urbanization. Catchment-scale data were compiled from remotely sensed datasets, whereas local-scale data were collected by citizen scientists as part of the global FreshWater Watch (FWW) program. The local-scale data included nearly 2600 coincident measurements of turbidity and observations of other local characteristics taken with a common method between March 2013 and June 2016 across 127 unique locations in 6 major population centers: Vancouver (Canada), São Paulo (Brazil), Curitiba (Brazil), Buenos Aires (Argentina), Hong Kong SAR (China), and Guangzhou-Foshan (China). Catchment- and site-scale information were modeled with Boosted Regression Trees (BRT) to identify likely drivers of increased turbidity both across the entire dataset and within individual cities. Urbanization was not consistently associated with turbidity. The global BRT model explained 60% of the variation in turbidity, and key predictors were catchment area, % of the catchment as grassland, rainfall, Gross Domestic Product, and % of the catchment as artificial surfaces. City-specific BRT models explained 35–67% of the variation in turbidity. Key predictors varied between cities and were often different than those observed at the global scale. Local-scale data collected by citizen scientists were less predictive of turbidity than catchment-scale factors and explained ~12% of the observed global variability in turbidity. Factors such as riverbank vegetation and the presence of point pollution sources explained some of the variation in turbidity, indicating their management could help mitigate elevated turbidity and sediment load in some urban rivers. Through this high-resolution, site-scale information, we highlight how community-sourced data may add value to freshwater monitoring programs.

**Key words:** NTU, turbidity, land use, catchment, water quality, urbanization, urban stream syndrome, boosted regression trees

Securing access to an adequate supply of quality freshwater is a major challenge in much of the world, especially in areas undergoing urbanization. Currently, 75% of the population in developed countries lives in urban areas (i.e., cities, towns, and suburbs), and 60% of the global population is expected to live in urban areas by 2030 (Paul and Meyer 2001, Cohen 2006, UNPD 2014). Continued urbanization represents a threat to freshwater ecosystems and major services, such as freshwater supply, flood mitigation, carbon storage, and soil fertility (Eigenbrod et al. 2011). The degradation of stream ecosystems caused by catchment urbanization is a recurring issue across the globe (Paul and Meyer 2001, Peters 2009, Walters et al. 2009), and has been referred to as the 'urban stream syndrome' (Walsh et al. 2005b). Major changes associated with this syndrome include: increases in the quantity and variety of pollutants in runoff, more erratic hydrology caused by increased impervious surface area and runoff conveyance, increased water

E-mail addresses: [4]leticia.m.ch@gmail.com; [5]eleanore.heasley@kcl.ac.uk; [6]sloiselle@earthwatch.org.uk; [7]i.thornhill@bathspa.ac.uk

DOI: 10.1086/703460. Received 14 July 2017; Accepted 13 October 2018; Published online 11 April 2019.
Freshwater Science. 2019. 38(2):303–320. © 2019 by The Society for Freshwater Science.

temperatures caused by riparian vegetation loss and warming of surface runoff on exposed de-vegetated artificial surfaces, reduction in channel and habitat complexity, channelization, and restricted interactions between the river and its land margin (Paul and Meyer 2001, Roy et al. 2003, Bernhardt and Palmer 2007, Vairavamoorthy et al. 2008, O'Driscoll et al. 2010). A deeper understanding of the relationship between stream characteristics and the type of urban development can provide valuable guidance for the management and restoration of urban streams and the associated downstream ecosystems (Meyer et al. 2005, Wenger et al. 2009).

In this study, we explore the effect of urbanization on stream water turbidity at both the catchment and site scales. Turbidity occurs when dissolved or particulate matter increases the scattering and absorption of visible radiation in water. This matter can be either organic, such as algae, or inorganic (Henley et al. 2000, Wenger et al. 2009). Substantial increases in turbidity may influence habitat quality, light penetration, nutrient dynamics, and algal and bacterial growth. Changes in local stream ecology at the individual, population, and community levels in response to increasing turbidity have also been reported (Ryan 1991, Henley et al. 2000, Walters et al. 2003, Paule-Mercado et al. 2016). Increased turbidity, relative to natural background values, has been associated with an increase of urban land cover. Examples include Georgia (USA; Roy et al. 2003), Seattle (USA; Brett et al. 2005), and South Korea (Chang 2008). Nevertheless, the effect of urbanization on turbidity remains unclear. Walsh et al. (2005b) reviewed the general symptoms associated with the urban stream syndrome and found that suspended solids (TSS) had an inconsistent response to urbanization. We expected to find a similar inconsistent relationship between urbanization and turbidity in this study, as turbidity is high correlated with TSS (Huey and Meyer 2010, Rügner et al. 2013, West and Scott 2016).

A major challenge in mitigating elevated turbidity in urban areas is the limited availability of local information. The collection of data through the FreshWater Watch (FWW) program (Loiselle et al. 2017) presents an excellent opportunity to study turbidity in urban areas that are in different geographic regions. Large-scale studies of water quality are rare because of methodological inconsistencies that limit data comparison. The FWW program avoids this problem as all participants follow a consistent methodology when they collect water quality data and record information on land and river conditions. Data collected by FWW participants have been used to study the drivers of eutrophication and algal blooms (Cunha et al. 2017, Zhang et al. 2017). In this study, we used FWW data to assess the relevance of data collected by citizen scientists in the study of turbidity and to identity which factors potentially influence turbidity across a wide range of urban settings. To do so, we used data from Vancouver (Canada), São Paulo (Brazil), Curitiba (Brazil), Buenos Aires (Argentina), Hong Kong SAR (China), and Guangzhou-Foshan (China).

## METHODS

We used boosted regression trees to model the relationship between stream water turbidity measured by FWW participants and local- and catchment-scale environmental conditions. Turbidity and other site-scale information (bank vegetation and land use) were gathered by FWW participants. In addition, rainfall was calculated at the site scale from spatial datasets. Catchment-scale information (land use, population density, Gross Domestic Product [GDP], and catchment morphologic characteristics) was extracted from available spatial datasets (Table 1). Of all the variables, proportion of artificial surfaces in the catchment (as part of land use), GDP, and population density are probably most closely associated with urbanization. To study the potential influence of urbanization and other factors on turbidity, we ran 1 model with all the sampled sites and variables (hereafter 'global model') and 1 model for each city. To assess the usefulness of the information collected by citizen scientists, we then ran 1 model with only turbidity and citizen science derived information.

### Site-scale information

Data used at the site-scale were turbidity, bank vegetation, and local land use that were recorded by FWW participating citizen scientists and rainfall obtained from a global database (Kanamitsu et al. 2002).

The FWW program (freshwaterwatch.thewaterhub.org) is an Earthwatch platform that helps citizen scientists study their local freshwater ecosystems with methods that are consistent across the globe. The platform was launched in 2012 and supports global and local monitoring by following a set of core global protocols, with parameters added locally as required. All participants are trained to collect data following common methods (Castilla et al. 2015; https://freshwaterwatch.thewaterhub.org), which they can upload directly to a global database with a smartphone application or website portal. Site selection, sampling frequency, and the collection of additional variables depend on the objectives of each local project. Sample collection in the FWW program was not always consistent because of the flexibility afforded to participants and local projects, which resulted in samples that were not evenly spread temporally or spatially. These differences could bias results toward types of sites or climatic conditions that are most common. To reduce this bias and make our model more robust, we selected data from sites with similar numbers of samples and with representation across the seasons.

In this study, we included 127 rivers and stream sites within 6 urban stream catchments that were sampled at least 1 time each season between March 2013 and June 2016 (Fig. 1). Study locations, with their respective number of sample sites, included Buenos Aires ($n$ = 18; Argentina), Curitiba ($n$ = 26; Brazil), Guangzhou ($n$ = 19; China), Hong Kong SAR ($n$ = 21; China), São Paulo ($n$ = 23; Brazil),

Table 1. List of candidate predictors used in the boosted regression tree modeling. GIS = Geographic Information System, Spatial data. CS = data collected by citizen scientists. FAO GeoNetwork = Food and Agriculture Organization of the United Nations (http://www.fao.org/geonetwork/srv/en/main.home). WWF = World Wildlife Fund (http://hydrosheds.org/). SEDAC GPW = Socioeconomic Data and Applications Center of NASA, Gridded Population of the World (http://sedac.ciesin.columbia.edu/data/collection/gpw-v4). Global Risk Data Platform, World Bank (https://preview.grid.unep.ch/). FWW = FreshWater Watch (https://freshwaterwatch.thewaterhub.org/). SEDAC NCEP = Socioeconomic Data and Applications Center of NASA, National Center for Environmental Prediction (http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html)

| | Variable | Unit/measure | Type of data | Source of information |
|---|---|---|---|---|
| Catchment Scale | Artificial surfaces | % coverage | GIS | FAO GeoNetwork |
| | Cropland | % coverage | GIS | FAO GeoNetwork |
| | Grassland | % coverage | GIS | FAO GeoNetwork |
| | Tree covered areas | % coverage | GIS | FAO GeoNetwork |
| | Shrub covered areas | % coverage | GIS | FAO GeoNetwork |
| | Catchment area | $km^2$ | GIS | WWF |
| | Catchment slope | % | GIS | WWF |
| | Population density | people/$km^2$ | GIS | SEDAC GPW-v4 |
| | Gross Domestic Product | thousands of constant 2000 USD | GIS | Global Risk Data Platform, World Bank |
| | Terrestrial ecoregions | categorical | GIS | WWF |
| Site Scale | Turbidity | NTU | CS | FWW |
| | Average 72 h rainfall | mm/h | GIS | SEDAC NCEP |
| | Riverbank bare ground | presence/absence | CS | FWW |
| | Riverbank grass | presence/absence | CS | FWW |
| | Riverbank trees and shrubs | presence/absence | CS | FWW |
| | Point source discharge | presence/absence | CS | FWW |
| | Urban residential | presence/absence | CS | FWW |
| | Urban park | presence/absence | CS | FWW |
| | Industrial | presence/absence | CS | FWW |
| | Agriculture | presence/absence | CS | FWW |
| | Forest | presence/absence | CS | FWW |
| | Grassland and shrubs | presence/absence | CS | FWW |
| | Season | categorical | – | – |

and Vancouver ($n = 20$; Canada). The combined dataset included 2565 coincident measurements of turbidity and observations of site-scale information (i.e., bank vegetation and land use). Geographic coordinates and measurement times were recorded for each sample (Tables S1–S6).

Citizen scientists used calibrated Secchi tubes to measure turbidity with detection limits between 12 and 240 NTU (Tyler 1968, Preisendorfer 1986, Wernand 2010). Secchi depth has been successfully used in citizen science programs (Lathrop et al. 1996, Bruhn and Soranno 2005, Lottig et al. 2014) to provide robust estimates of turbidity that are comparable to scatterometer, turbidity, and suspended particulate measurements taken by professional scientists (Obrecht et al. 1998, Canfield et al. 2002, Castilla et al. 2015). In this study, a significant number of observations were below or above detection limits. Consequently, turbidity measurements equal to, or below the lower detection limit (44% of samples) were recorded as 12 NTU, and those equal to or above the upper detection limit (3% of samples) were recorded as 240 NTU. The lower detection limit of the Secchi tube is a function of its length. To achieve a 5 NTU detection limit (a typical value for perceived turbidity in natural waters, Myre and Shaw 2006), the Secchi tube would need to be nearly 2× its actual length of 48 cm. The use of such long tubes would have further complicated measurements and increased costs of this global project. The 48-cm long tube was a compromise between costs and detection limits, which limited the interpretability of turbidity levels in rivers with low turbidity.

At each site, citizen scientists also recorded information about presence/absence of riverbank cover (trees/shrubs, grass, bare ground), presence of point discharge sources of pollution, and dominant land use in the immediate vicinity of the sampling site (i.e., urban residential, urban park, industrial, agriculture, forest, and grassland and shrubs) (Table 1). These measurements were based on a reach length

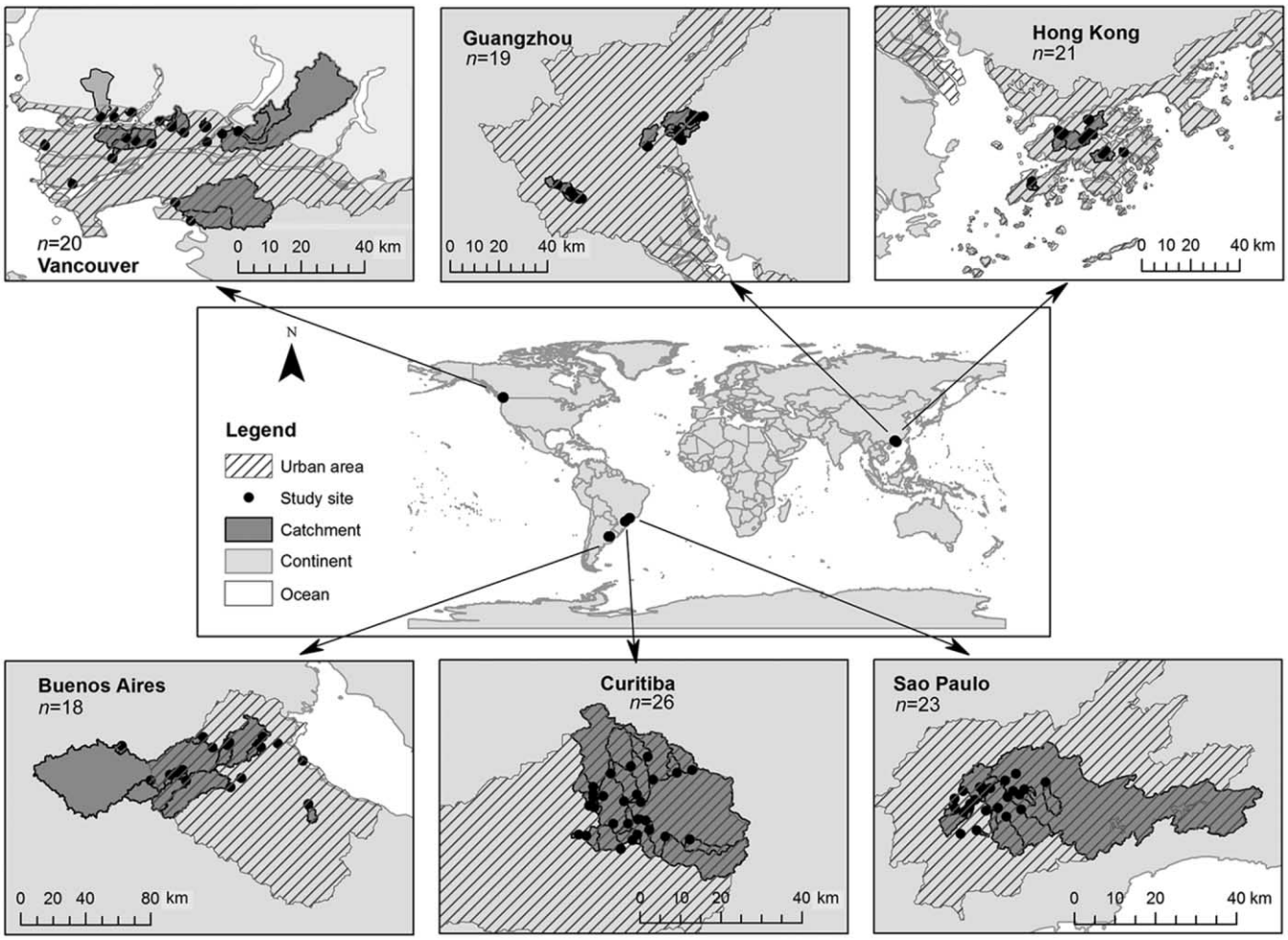


Figure 1. Study sites, urban areas, and their global context. The black circles in the center rectangle of this figure represent each of the 6 regional study locations. The black circles in the outside rectangle represent each site within each city. The number of sites in each city is indicated by an *n* in the figure. Some local catchments boundaries are hidden because of their close proximity to others or because they are too small to be visible at the map scale.

of 50 m upstream from the sampled site, widths of 50 m on each side of the river for riverbank characteristics and pollution sources, and widths of 100 m for land use.

We attributed each measurement to a season (spring, summer, autumn, and winter) as a categorical variable to assess the effects of climatic variability. Average rainfall was calculated for the 72 h period prior to each sampling event with data from the Physical Sciences Division of the National Oceanic and Atmospheric Administration (Kanamitsu et al. 2002). We used the *RNCEP* package in R, following the methods described in Kemp et al. (2012), to average the rainfall data at a spatial resolution of 30 arc-seconds (~1 km).

## Catchment-scale information

In addition to the information collected at the site scale, we obtained catchment-scale spatial information such as land use, demographic variables, climate variables, and catchment morphology characteristics.

We used the *Watershed Tool* in the *Hydrology Toolbox* in ArcMap 10.4.1 (ESRI, Redlands, California) to calculate flow directions from 3 arc-second (~100 m) resolution WWF-HydroSHEDS (hydrosheds.org) maps (Lehner et al. 2008). We then used the *Zonal Statistic as Table* function within the Spatial Analyst Tools in ArcMap to calculate land use, demographic variables, and ecoregions within each catchment.

We used the Global Land Cover SHARE (GLC-SHARE) database to obtain land cover information. GLC-SHARE is a land cover database created by the Land and Water Division of the Food and Agriculture Organization of the United Nations (FAO) that combines the best available high resolution national, sub-national, and regional information to produce a global map with a resolution of 30 arc-seconds (Latham et al. 2014). For China and Canada, the FAO used

regional base information to create GLC-SHARE coverages (Landsat 2002 for China, and MODIS and FAO LCCS2005 for Canada), whereas for Brazil and Argentina, the FAO used global information (GlobCover 2009, MODIS VCF 2010, Cropland database 2012). The GLC-SHARE produces a unified classification of the spatial information that includes the following categories: artificial surfaces, cropland, grassland, tree-covered area, shrub-covered area, herbaceous vegetation, mangroves, sparse vegetation, bare soil, water bodies, and snow and glacier cover.

We selected 2 demographic variables, population density and GDP as they are often used as indicators of the type of urban development present and the intensity of economic activities in urban areas. Population density was obtained from the Gridded Population of the World (GPWv4; CIESIN 2016). We used the GDP information from 2010 provided by The World Bank (available at preview.grid.unep.ch). Both datasets have a spatial resolution of 30 arc-seconds (~1 km).

We used the terrestrial ecoregions defined by Olson et al. (2001) (worldwildlife.org/publications/terrestrial-ecoregions-of-the-world) to assess the effects of regional differences in climate on turbidity.

Finally, we included catchment area and average catchment slope variables to represent catchment morphology. Catchment area was extracted from the attribute table of the catchments delineated in ArcMap. The slope for each catchment was calculated with the *Slope Tool* in Spatial Analyst Tools (ArcMap 10.4.1) based on 3 arc-second (~100 m) resolution digital elevation model, WWF-HydroSHEDS (hydrosheds.org).

## Data analyses

We used individual (not averaged) measurements of turbidity and associated site- and catchment-scale observations in all models.

We used boosted regression trees (BRT) to identify associations between turbidity and urbanization variables. By fitting BRT, we were also able to identify which citizen scientist acquired data were most strongly associated with turbidity. BRT combine the strengths of regression trees (models that relate a response to their predictors by recursive binary splits) and boosting (an adaptive method for combining many simple models to improve predictive performance) (Friedman 2001, Elith et al. 2008). In BRT, simple regression trees are fit iteratively to produce a single model. The first regression tree is selected to minimize the loss of information, and in each following step, model accuracy is improved by fitting a tree that reduces the residual deviance (error in the prediction) of the previous tree. The final BRT model is an additive regression model in which individual terms are trees (Elith et al. 2008). We created BRT models with the *gbm* package (version 2.1.1; Ridgeway 2015) in R 3.0.1 (R Core Development Team 2018).

BRT models require 3 parameters to be set: learning rate, bag fraction, and tree complexity (Elith et al. 2008). Learning rate, or shrinkage parameter, determines the contribution of each tree to the growing model. Slower learning rates result in better predictions, but this demand must be balanced with computing resources and time. Learning rate was set as 0.01 for the global model, and 0.001 for the models at city-scales and the model with citizen scientist data. Bag fraction is the proportion of data used at each step to fit a tree model. The bag fraction is extracted at random from the dataset. The addition of stochasticity improves the accuracy and speed of boosted models and reduces overfitting (Friedman 2002, Elith et al. 2008). In the global model the bag fraction was 0.6 (i.e., 60% of the data), whereas the city-scale and the citizen scientist model had bag fractions of 0.7 (i.e., 70%). The increase in the bag fraction in the city and citizen scientist models is because of the relatively smaller datasets. Tree complexity refers to the maximum number of nodes in a tree and limits the maximum number of interactions among predictors (Elith et al. 2008). This parameter was set at 5 for all models.

BRT, like many other statistical approaches, are vulnerable to model overfitting when input variables are highly correlated (Olden et al. 2008). Therefore, before introducing the predictors in the model, we used Spearman correlations to determine the correlations between variables (Tables S7–S13). When predictors were correlated ($rs > 0.7$), we only retained 1 of the correlated predictors in the model (Dormann et al. 2013). When possible, we retained the same variables for all the models to facilitate the discussion of the results. We calculated correlations with the *rcorr* function in the R *Hmisc* package (version 3.15; Harrel and Dupont 2015). We log-transformed and then standardized continuous predictor variables as that improved the interpretation of the BRT models. We used the *rescale* formula in the R *scales* package (version 2.1.3; Ridgeway 2015).

We fit the models with the function *gbm.step* in the R *gbm* package. This function fits a BRT model with cross-validation to estimate the optimal number of trees. After fitting the model, we used k-fold cross validation with *gbm.simplify* in the R *gbm* package to assess whether there were non-informative predictors that could be removed to improve the model (Elith et al. 2008). The relative importance of each BRT model parameter was scaled so that the sum of all parameters was 100. A higher percentage of a variable indicates stronger relative importance on the response. Variables with a smaller model importance than that expected by chance (100% divided by the number of variables) were considered irrelevant for interpretation (Müller et al. 2013). We created partial dependence plots for each model variable to describe the relationship between predictors and turbidity. Partial dependence plots show the effect of a variable on turbidity (the variable response) after accounting for the average effects of all other variables in the model (marginal ef-

fect) (Pearson 2017). These plots provide a visual basis for interpretation of results.

## RESULTS

### Data overview

Here, we briefly describe patterns of variation in turbidity and the predictor variables at the catchment- and site-scales across the dataset as a whole and by city (Figs 2A–J, 3A–K). Turbidity measurements covered the full range of detection at both global- and city-scales (i.e., 12–240 NTU). Globally, turbidity averaged 38 NTU, with a large difference between cities. For example, turbidity was highest in São Paulo (mean = 77 NTU), and lowest in Vancouver (mean = 16 NTU) (Fig. 2A).

Globally, the studied catchments were covered predominantly by trees (range 0–93%, mean = 30%; Fig. 2E) and artificial surfaces (range 0–100%, mean = 29%; Fig. 2B),

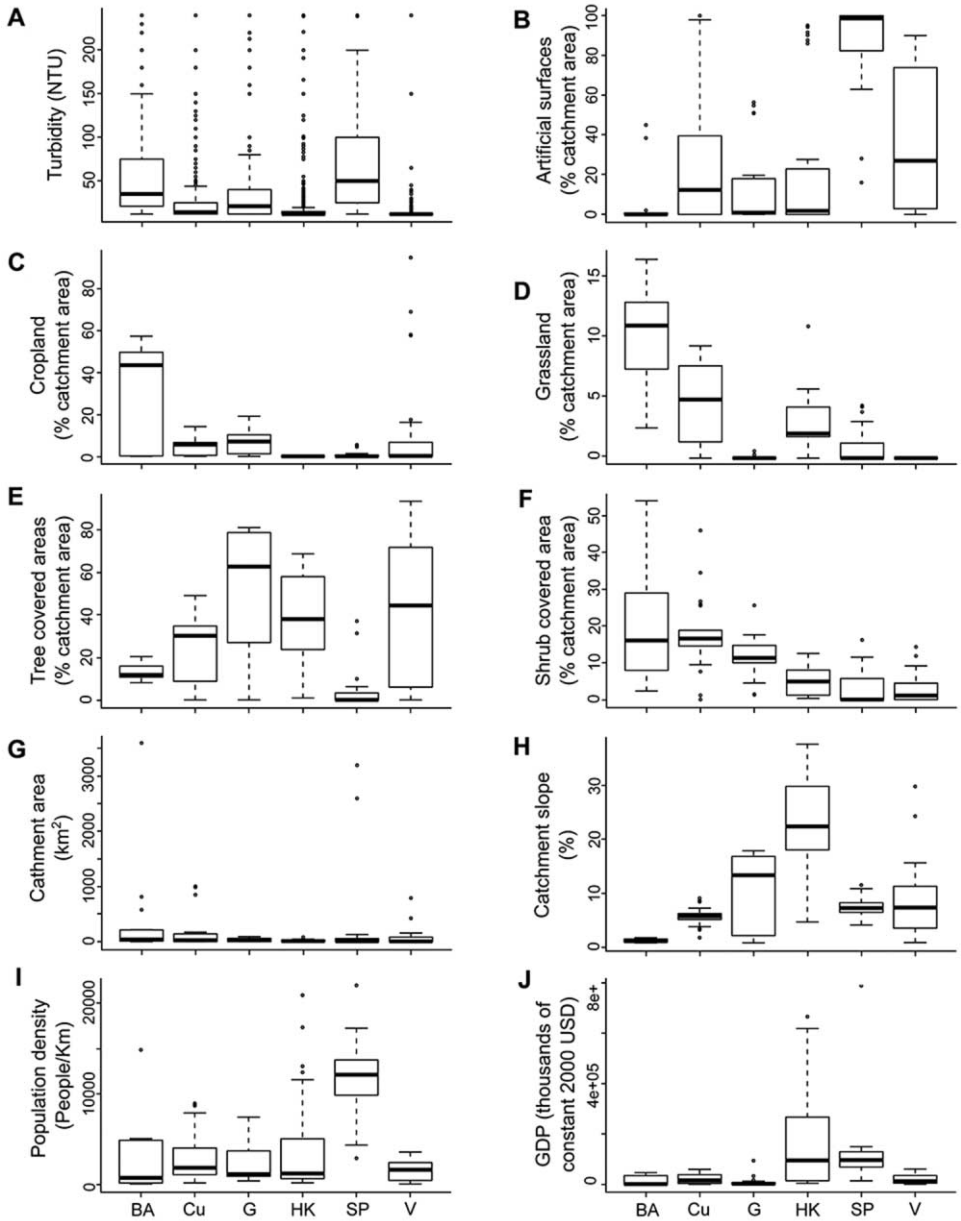


Figure 2. Boxplots showing variation among sites within each city in turbidity and 9 catchment-scale predictors: turbidity (A), artificial surfaces (B), cropland (C), grassland (D), tree covered areas (E), shrub covered area (F), catchment area (G), catchment slope (H), population density (I), GDP (J). Horizontal bars represent the means; the box ends indicate the 25th and 75th percentiles; whiskers indicate the 10th and 90th percentiles; circles indicate outliers. BA = Buenos Aires, Cu = Curitiba, G = Guangzhou, HK = Hong Kong, SP = São Paulo, and V = Vancouver, GDP = gross domestic product.

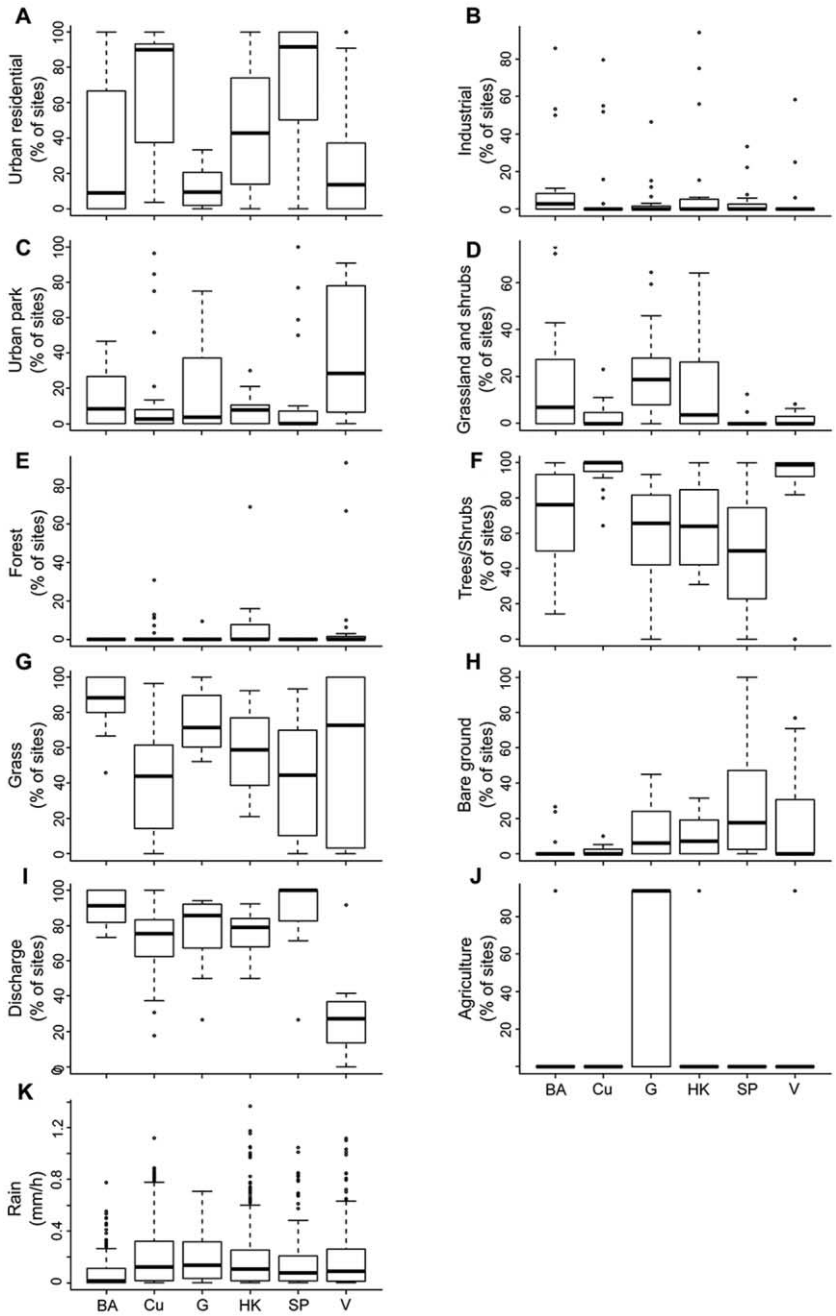


Figure 3. Boxplots showing variation among sites within each city in 11 site-scale predictors: urban residential (A), industrial (B), urban park (C), grassland and shrubs (D), forest (E), trees/shrubs (riverbank trees and shrubs) (F), grass (riverbank grass) (G), bare ground (riverbank bare ground) (H), discharge (point source discharge) (I), agriculture (J), and rain (average 72 h rainfall) (K). '% of sites' refers to the percentage of sites with the attribute present. Horizontal bars represent the means; the box ends indicate the 25$^{th}$ and 75$^{th}$ percentiles; whiskers indicate the 10$^{th}$ and 90$^{th}$ percentiles; circles indicate outliers.

with the exception of Buenos Aires, where the dominant land use was cropland (range 0–57%, mean = 29%; Fig. 2C). Grassland was least abundant across study sites (range 0–17%, mean = 3%; Fig. 2D). Catchment size also varied considerably among cities. Buenos Aires and São Paulo had the largest catchments in this study and were on average >300 km$^2$(Buenos Aires: range 0.5–3600 km$^2$, São Paulo: range 0.5–3200 km$^2$), whereas Hong Kong and Guangzhou had the smallest catchments averaging <50 km$^2$ (Hong Kong: range 0.6–81 km$^2$, Guangzhou: 0.1–92 km$^2$) (Fig. 2G). Catchment slope ranged from a mean of 1.3% in Buenos Aires (range 0.9–1.9%) to a mean of 23% in Hong Kong (range 4.8–37.6%) (Fig. 2H). Average population density was 4000 people/km$^2$ in the studied cities (range 20–22,000 people/

km$^2$), with maximum values in São Paulo (average: 12,000 people/km$^2$, range: 2900–22,000 people/km$^2$) (Fig. 2I). GDP (in US dollars set to the year 2000) was highest in Hong Kong (average 170,500; range 4200–666,000) and São Paulo (average 118,750; range 13,840–789,400), and lowest in Guangzhou (average 7500; range 295–94,000) (Fig. 2J). Curitiba, Guangzhou, Hong Kong, and São Paulo are located in the 'tropical/subtropical moist broadleaf forest' ecoregion, Vancouver is in the 'temperate coniferous forest' ecoregion, and Buenos Aires is in the 'temperate grasslands, savannas, shrublands' ecoregion.

Most of the study sites were located in residential areas (45% of the sites; Fig. 3A). However, in Guangzhou, most of the sampling sites (64%) were located in catchments with substantial agricultural land cover (Fig. 3J). Approximately 75% of all study site riverbanks were tree covered (Fig. 3F), and 72% had potential point source discharges nearby (Fig. 3I). These values were similar among cities, with the exception of São Paulo, where only 50% of sites had trees on the riverbank (Fig. 3F), and Vancouver, where only 32% of sites had potential point source discharges nearby (Fig. 3I). Precipitation within 72 h prior to sampling was low, averaging 0.2 mm/h (range 0–1.37 mm/h) at both the global and city scales, except for Buenos Aires, which had values of 0.07 mm/h (range 0–0.78 mm/h) (Fig. 3K).

## Global correlates of turbidity

The global model included 12 predictors and explained 62% of the variability in turbidity (Table 2). Optimal model fit was achieved at 2400 trees. City, catchment area, grassland, average 72-h rainfall, GDP, and percentage of artificial surfaces were important explanatory variables following the criteria of Müller et al. (2013) (Table 2). Partial plots showed that cities had different turbidity levels (Fig. 4A), and that turbidity was higher in larger catchments and in those with a higher proportion of grassland (Table 2; Fig. 4B, C). Turbidity also increased with rainfall, but it plateaued at higher values of precipitation (Table 2; Fig. 4D). We detected a bimodal response of turbidity to GDP (Fig. 4E) and artificial surfaces, where turbidity (Fig. 4F) was highest in both low- and high-income areas and in catchments with either high or low artificial surface coverage (Table 2).

## City-scale correlates of turbidity

***São Paulo*** The final model for São Paulo included 11 variables and explained 67% of the variability in turbidity. This model fit was achieved at 6000 trees (Table 2). The variables population density, mean catchment slope, average 72-h rainfall, GDP, and cropland catchment cover explained most of the variability. Areas with lower population density had higher turbidity (Fig. 5A), as did catchments with a moderate slope compared with either low or high slope (Fig. 5B). The association of turbidity with rainfall in this model was similar to that in the global model with higher values of turbidity at lower rain levels (Fig. 5C). Turbidity increased with GDP (Fig. 5D) and was higher in catchments where cropland cover was low (Fig. 5E).

***Hong Kong*** The final model for Hong Kong included 11 variables and explained 66% of the variation in turbidity (Table 2). This model fit was achieved at 4350 trees. Grassland, artificial surface catchment cover, and average 72-h rainfall were the most important variables. Turbidity increased with the percentage of grass and artificial surfaces in the catchment (Fig. 6A, B) and decreased marginally with increased rainfall (Fig. 6C).

***Guangzhou*** The final model for Guangzhou included 13 variables and explained 56% of the variability in turbidity (Table 2). This model fit was achieved at 6500 trees. Average 72-h rainfall, catchment area, artificial surface catchment cover, season, and catchment cropland cover were the most important variables. Turbidity increased with rainfall (Fig. 7A), with increasing catchment size (Fig. 7B), and with a high coverage of artificial surfaces (Fig. 7C). Summer and autumn had the highest values of turbidity (Fig. 7D). Turbidity was higher in catchments with relatively low cropland catchment coverage, but this relationship was not linear because there was a peak in turbidity at high cropland coverage (Fig. 7E).

***Buenos Aires*** The final model for Buenos Aires included 14 variables and explained 50% of the variability in turbidity (Table 2). This model fit was achieved at 3400 trees. Catchment area, average 72-h rainfall, grassland and cropland catchment cover, and season explained most of the variability in turbidity. Large catchments and catchments high grassland coverage had the highest turbidity (Fig. 8A, C), whereas catchments with low crop coverage had the lowest turbidity (Fig. 8D). Elevated precipitation was generally associated with high turbidity (Fig. 8B), and turbidity was lowest in the summer (Fig. 8E).

***Curitiba*** The model for Curitiba included 13 variables and explained 50% of the variability in turbidity (Table 2). This model fit was achieved at 4150 trees. Average 72-h rainfall, GDP, catchment area, and shrub and artificial surface catchment cover explained most of the variability in turbidity. Turbidity increased with higher rainfall (Fig. 9A) in catchments with a higher GDP (Fig. 9B) and in catchments with a lower proportion of artificial surfaces (Fig. 9E). Turbidity values also increased with increasing catchment size (Fig. 9C). Shrub cover was not strongly related to turbidity as values were relatively constant across low- to mid-values of shrub cover. A higher values of shrub cover, turbidity initially decreased slightly with increasing shrub cover and then increased with further increases in shrub cover (Fig. 9D). However, this trend was of limited importance

Table 2. Results of the BRT model at the global- and city- scales, with site- (including citizen scientist data) and catchment-level data. Relative variable importance is shown in bold when it is greater than expected by chance (after Müller et al. 2013). The symbols next to relative variable importance approximate the marginal effects of the variable on turbidity based on partial plot interpretation (see Figs 4–9), where U = bimodal, ∩ = unimodal, / = positive, \ = negative and – = unclear. GDP = gross domestic product.

| | | Global | Hong Kong | São Paulo | Guangzhou | Vancouver | Buenos Aires | Curitiba |
|---|---|---|---|---|---|---|---|---|
| Overall variance explained by the model (%) | | 61.6 | 66.4 | 66.9 | 55.6 | 51.8 | 49.5 | 36.3 |
| Number of variables | | 13 | 11 | 11 | 13 | 13 | 14 | 13 |
| Tree complexity | | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Learning rate | | 0.01 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| City | | **30.9** | n/a | n/a | n/a | n/a | n/a | n/a |
| Catchment information (percentage of the model explained by each variable) | Catchment area | **10.5** / | **2.9** | | **20.4** / | 2.2 | **23.5** / | **12.8** / |
| | Grassland | **10.4** / | **42.8** / | | 1.4 | | **14.0** / | |
| | GDP | **9.9** U | | **14.2** / | | | | **17.1** / |
| | Artificial surfaces | **8.3** / | **33.0** / | 6.8 | **17.9** \ | **8.51** / | 2.7 | **7.8** \ |
| | Shrub covered areas | 5.9 | | | | 6.84 | | **9.1** \ |
| | Cropland | 4.8 | | **9.3** \ | **9.3** \ | **29.6** / | **10.2** \ | 5.8 |
| | Population density | | | **21.5** U | | | | |
| | Catchment slope | | | **20.5** ∩ | | | | |
| | Tree covered areas | | | | 6 | 10.5 | | |
| Site information (percentage of the model explained by each driver) | Average 72 h rainfall | **10.3** ∩ | **11.4** \ | **14.6** \ | **23.9** ∩ | **25.7 U** | **20.4** / | **32.1** ∩ |
| | Season | 3.4 | 5.1 | 4.6 | **15.8** ∩ | **8.95** – | **10.0** – | 4.89 |
| | Urban residential | 2.1 | 0.7 | 1.2 | | 2 | 1.5 | 3.3 |
| | Riverbank trees and shrubs | 1.2 | 1.5 | | 5 | | 1.2 | 1.7 |
| | Point source discharge | 1.2 | 1.8 | 2.8 | 1.9 | 3.37 | 2 | 0.2 |
| | Riverbank grass | 1.2 | 0.6 | 2.9 | 1.7 | 2.19 | 3.2 | 1.7 |
| | Agriculture | | 0 | | 1.1 | 0.1 | 2.5 | |
| | Riverbank bare ground | | | 1.7 | 0.5 | 0 | 0 | 0 |
| | Grassland and shrub | | 0.3 | | 0.6 | | | |
| | Urban Park | | | | 0.5 | | 0.9 | 3.6 |
| | Industrial | | | | | | | |
| | Forest | | | | | | | 0.0 |

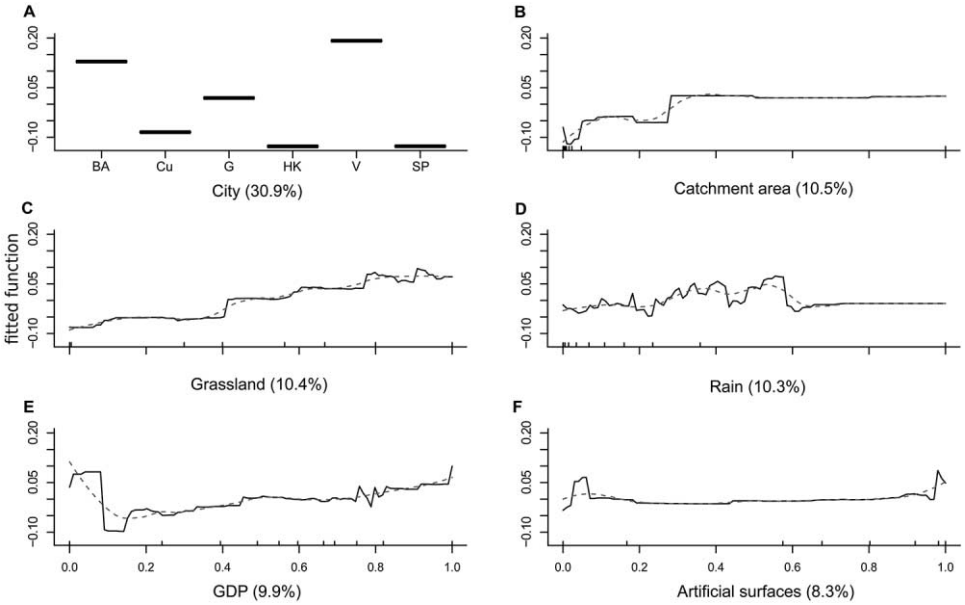


Figure 4. Partial plots for the boosted regression tree (BRT) models of turbidity at the global scale. Plots are ordered by importance of the covariates in the final boosted regression tree model (Table 2): city (categorical variable) (A), catchment area (catchment-scale) (B), grassland (catchment-scale) (C), rain (average 72 h rainfall, site-scale) (D), GDP (gross domestic product, catchment-scale) (E), and artificial surfaces (catchment-scale) (F). Dashed lines show smoothed lines of best fit. Tick marks in the inner side of the *x*-axes of the plots (B–F) indicate the deciles (10% quantiles) of the observed distribution of continuous predictor variables.

given the small amount of variability in turbidity explained by shrub cover.

***Vancouver*** The model for Vancouver included 13 variables and explained 52% of the variability in turbidity (Table 2). This model fit was achieved at 2850 trees. Cropland, average 72-h rainfall, tree cover, season, and artificial surfaces explained most of the variability in turbidity. Turbidity increased with higher coverage of cropland (Fig. 10A) and also in catchments with higher proportions of artificial surfaces (Fig. 10D). Turbidity values were higher at the lowest and highest intensities of precipitation (Fig. 10B). Catch-

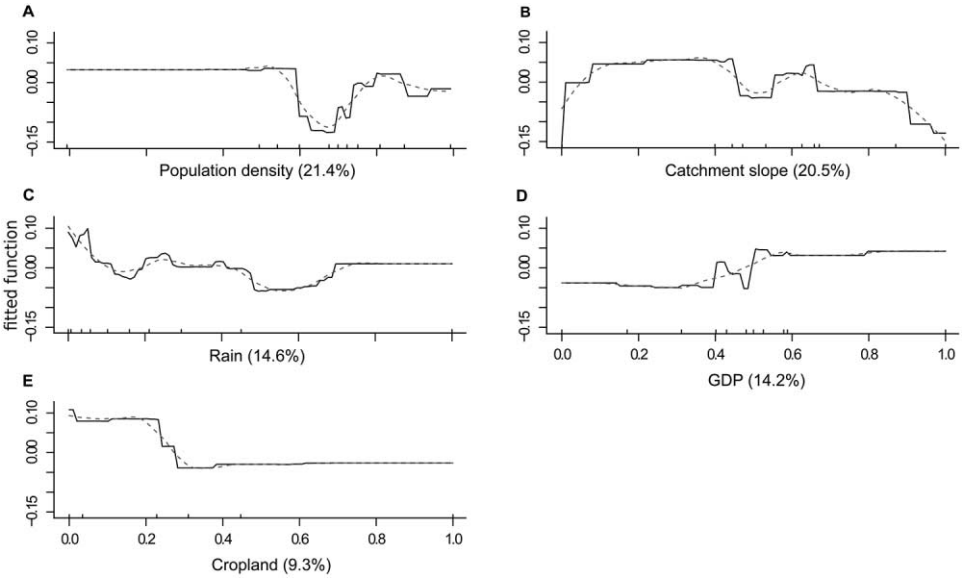


Figure 5. Partial plots for the boosted regression trees (BRT) model for turbidity in São Paulo. Plots are ordered by importance of the covariates in the final boosted regression tree model (Table 2): population density (catchment-scale) (A), catchment slope (catchment-scale) (B), rain (average 72 h rainfall, site-scale) (C), GDP (gross domestic product, catchment-scale) (D), cropland (catchment-scale) (E). Dashed lines show smoothed lines of best fit. Tick marks in the inner side of the *x*-axes of the plots indicate the deciles (10% quantiles) of the observed distribution of continuous predictor variables.

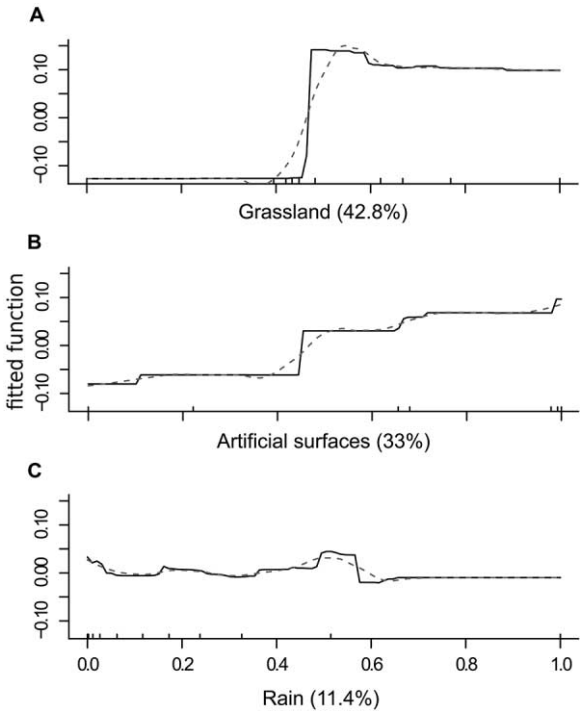


Figure 6. Partial plots for the boosted regression tree (BRT) models of turbidity in Hong Kong. Plots are ordered by importance of the covariates in the final boosted regression tree model (Table 2): grassland (catchment-scale) (A), artificial surfaces (catchment-scale) (B), rain (average 72 h rainfall, site-scale) (C). Dashed lines show smoothed lines of best fit. Tick marks in the inner side of the *x*-axes of the plots indicate the deciles (10% quantiles) of the observed distribution of continuous predictor variables.

ments with higher proportion of tree coverage were associated with high turbidity (Fig. 10C) and turbidity was highest in winter (Fig. 10E).

### Contribution of citizen scientist derived information

This global model was derived only from data collected by FWW citizen scientists, included the 9 site scale binary variables, and explained 11.6% of the variability in turbidity (Table 3). The model was fitted with 3950 trees. Turbidity tended to be higher where potential point source discharges were identified (Table 3; Fig. 11A). In contrast, turbidity was lower when riverbank vegetation included trees, shrubs, or both, and when urban residential land use was the primary land cover (Table 3; Fig. 11B, C). No models could be fit at the city scale based on citizen scientist data alone.

## DISCUSSION

Our findings suggest that turbidity did not uniformly follow the expected patterns of the urban stream syndrome (i.e., decrease of water quality with increased urbanization), confirming our initial hypothesis of an inconsistent relationship between turbidity and urbanization. Factors associated with urbanization were not the main drivers of turbidity at a global scale, nor were they consistently associated with turbidity in the city-scale models. We also found that information collected by citizen scientists could contribute to understanding of the drivers of turbidity and, thus, provide information that may help inform management in urban areas.

### Turbidity and the urban stream syndrome

The main results of this study were that site location (i.e., city) is a highly important explanatory variable for turbidity, and that there is a lack of consistency in the factors that predict turbidity at this scale. The urban stream syndrome describes how streams in urbanized catchments are ecologically degraded (Paul and Meyer 2001, Meyer et al. 2005, Walsh et al. 2005a). Walsh et al. (2005b) reviewed the symptoms of the urban stream syndrome looking for consistent trends across geographic regions. In their review TSS, a variable strongly correlated with turbidity (Huey and Meyer 2010, Rügner et al. 2013, West and Scott 2016), responded inconsistently to urbanization (Walsh et al. 2005b). Similarly, in this study we found that turbidity did not show a consistent geographic response in urban areas. Our results, therefore, support those of recent studies that have shown that stream response to urbanization can vary greatly among regions (Wenger et al. 2009, Booth et al. 2016, Capps et al. 2016). It should be noted that a significant proportion of the turbidity measurements were below the detection limit (12 NTU), complicating the interpretation of the lower end of the partial plot.

Among the factors associated with urbanization in this analysis (artificial surfaces, GDP, and population density), at least 1 significantly explained turbidity dynamics in most city-scale models as well as the global model. Nevertheless, urbanization factors explained less of the variability in turbidity than catchment area, proportion of grassland cover, and rainfall. This result was not surprising as sediment load depends not only on land use, but also on climatic conditions and soil characteristics (Hunsaker and Levine 1995). Geology can also have an important effect on turbidity of water bodies (Rains et al. 2008), therefore, future studies on turbidity would benefit from having an appropriate catchment indicator of geologic erodibility.

Our study showed that the proportion of artificial surfaces was associated with turbidity at global and city scales, although models differed in the relationship between these variables. Catchment imperviousness has been highlighted as one of the most important factors related to urban stream degradation (Paul and Meyer 2001, Walsh et al. 2005a, b, Wenger et al. 2009). An increase in artificial surfaces decreases the permeability of the catchment, which increases surface runoff (Dunne and Leopold 1978, Booth and Jackson 1997, Wenger et al. 2009). Increased runoff, in turn, can increase turbidity by increasing transport of sediments into waterways. In our study, the global model and the models

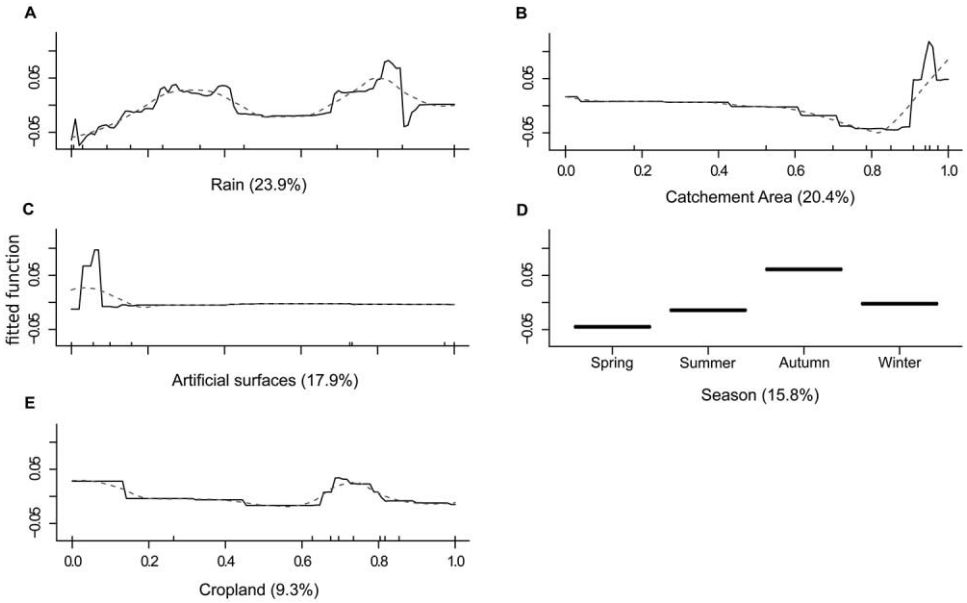


Figure 7. Partial plots for the boosted regression trees (BRT) model for turbidity in Guangzhou. Plots are ordered by importance of the covariates in the final boosted regression tree model (Table 2): rain (average 72 h rainfall, site-scale) (A), catchment area, (catchment-scale) (B), artificial surfaces (catchment-scale) (C), season (categorical) (D), cropland (catchment-scale) (E). Dashed lines show smoothed lines of best fit. Tick marks in the inner side of the *x*-axes of the plots (A–C, and E) indicate the deciles (10% quantiles) of the observed distribution of continuous predictor variables.

for Hong Kong and Vancouver showed this pattern; however, opposite trends were observed in the models for Guangzhou and Curitiba. A decrease in turbidity may occur when increases in impervious surfaces cause corresponding reductions in the bare land area, which is more susceptible to erosion (Moreno Madriñán et al. 2012).

On a global scale, turbidity was highest at both the lowest and highest values of GDP, creating a U-shape relationship. This association of turbidity with GDP could be related to a transition from farming catchments to very industrialized catchments, such that the lowest values of turbidity occur at intermediate levels of urbanization. The models for

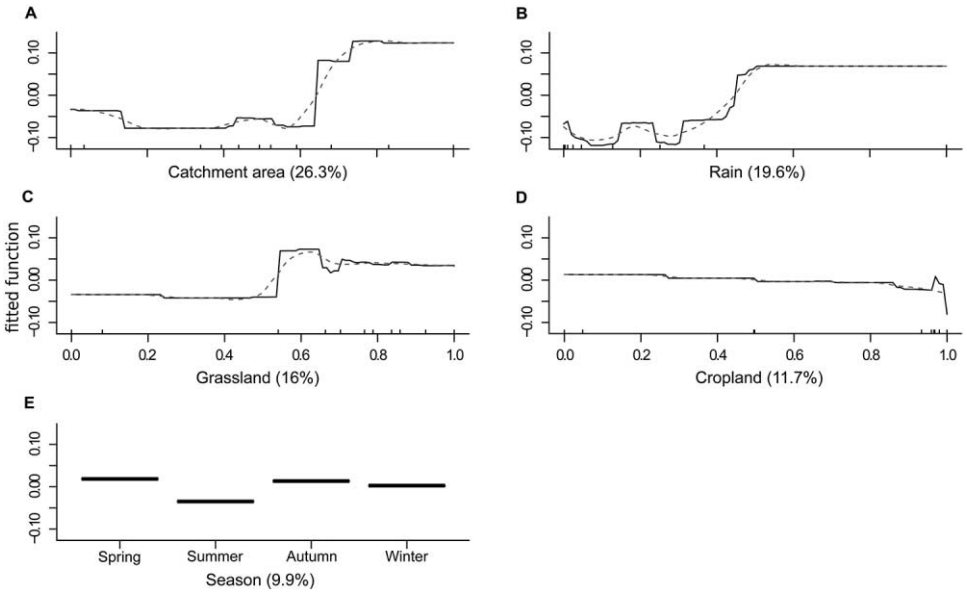


Figure 8. Partial plots for the boosted regression trees (BRT) model for turbidity in Buenos Aires. Plots are ordered by importance of the covariates in the final boosted regression tree model (Table 2): catchment area (catchment-scale) (A), rain (average 72 h rainfall, site-scale) (B), grassland (catchment-scale) (C), cropland (catchment-scale) (D), season (categorical) (E). Dashed lines show smoothed lines of best fit. Tick marks in the inner side of the *x*-axes of the plots (A–D) indicate the deciles (10% quantiles) of the observed distribution of continuous predictor variables.

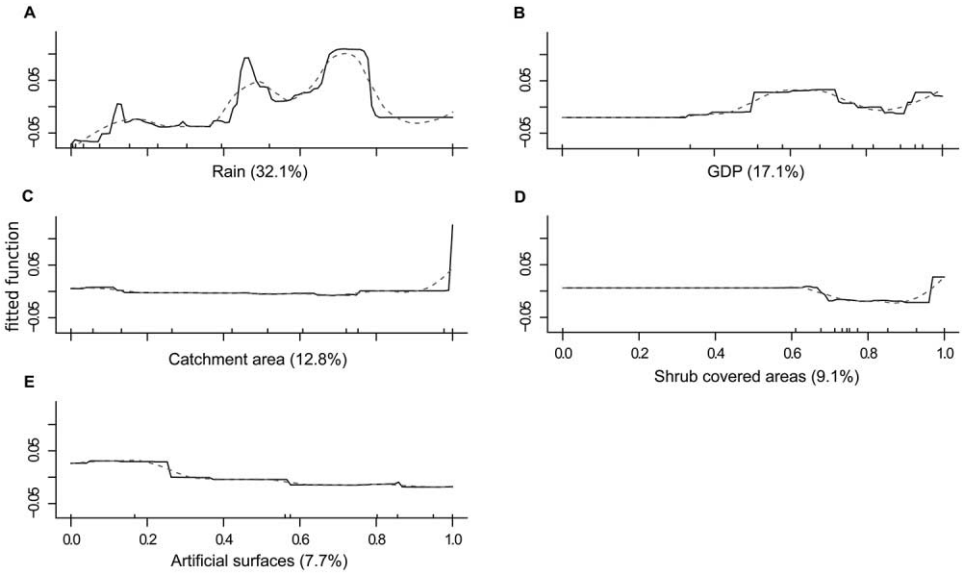


Figure 9. Partial plots for the boosted regression trees (BRT) model for turbidity in Curitiba. Plots are ordered by importance of the covariates in the final boosted regression tree model (Table 2): rain (average 72 h rainfall, site-scale) (A), GDP (gross domestic product, catchment-scale) (B), catchment area (catchment-scale) (C), shrub covered areas (catchment-scale) (D), artificial surfaces (catchment-scale) (E). Dashed lines show smoothed lines of best fit. Tick marks in the inner side of the *x*-axes of the plots indicate the deciles (10% quantiles) of the observed distribution of continuous predictor variables.

São Paulo and Curitiba showed that catchments with higher GDP had higher values of turbidity. Similarly, Li et al. (2008) reported an increase of suspended particulate matter with economic growth and population density in the Han River basin. The decrease in water quality with the increase in GDP could be related to an increase in industrial or mining activities (Zhang et al. 2009, Zhou et al. 2012).

Population density was a significant predictor of turbidity only in São Paulo. In this city, turbidity was highest at both the lowest and highest values of population density, creating a U-shape relationship. The high levels of turbidity in areas of low population density could be related to interactions between land use types and the topography of the landscape. For example, correlations showed that popula-

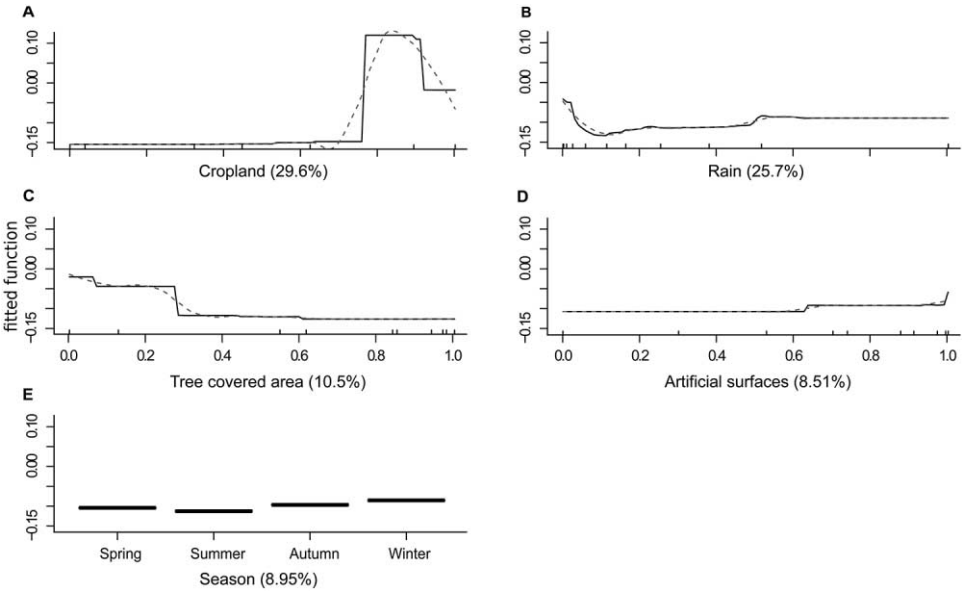


Figure 10. Partial plots for the boosted regression trees (BRT) model for turbidity in Vancouver. Plots are ordered by importance of the covariates in the final boosted regression tree model (Table 2): cropland (catchment-scale) (A), rain (average 72 h rainfall, site-scale) (B), tree covered area (C), artificial surfaces (catchment-scale) (D), season (categorical) (E). Dashed lines show smoothed lines of best fit. Tick marks in the inner side of the *x*-axes of the plots (A–D) indicate the deciles (10% quantiles) of the observed distribution of continuous predictor variables.

Table 3. Results of the global-scale BRT model, with only the site information collected by citizen scientists. Relative variable importance is shown in bold when it is greater than expected by chance (after Müller et al. 2013). The symbols approximate the marginal effects of the variable on turbidity based on partial plot interpretation (Fig. 10), where / = positive and \ = negative.

| | | Citizen scientist model |
|---|---|---|
| Overall variance explained by the model (%) | | 11.57 |
| Number of variables | | 9 |
| Tree complexity | | 5 |
| Learning rate | | 0.001 |
| Site information (percentage of the model explained by each driver) | Point source discharge | **26.0** / |
| | Riverbank trees and shrubs | **24.6** \ |
| | Urban residential | **13.3** \ |
| | Forest | 8.6 |
| | Agriculture | 8.0 |
| | Riverbank grass | 6.3 |
| | Urban park | 6.1 |
| | Riverbank bare ground | 5.1 |
| | Grassland and shrubs | 2.1 |

tion density in São Paulo increased with the proportion of artificial surfaces and decreased when the proportion of cropland, grassland, and shrubland and trees. An increase in turbidity with an increase in population is consistent with other studies that found population density was positively correlated with TSS (Huey and Meyer 2010, Rügner et al. 2013, West and Scott 2016).

## Catchment-scale predictors of turbidity

After the sampling site location (i.e., city, explained in the previous section) the factors most strongly associated with most of the variation in turbidity were catchment area, percent grassland cover, average 72-h rainfall, and cropland.

Catchment area was a significant predictor of turbidity in the global model and in the city-scale models of Guangzhou and Buenos Aires. This observation is consistent with previous studies that demonstrate catchment area can have a significant effect on stream water quality (Davies et al. 2000, Buck et al. 2004, Lawler et al. 2006b). In larger catchments, an increased sediment load to the lower reaches results from an increased number of sources and transport, resulting in an expected increase in turbidity.

Percent grassland cover in a catchment was a significant predictor of turbidity in the global, Hong Kong, and Buenos Aires models. Contrary to our expectations, higher grassland coverage was associated with higher values of turbidity in these catchments. In a global-scale study, García-Ruiz et al. (2015) found that high percentages of grassland cover had lower rates of erosion, with an expected reduction in turbidity. Similarly, Panagos et al. (2015) suggested that the conversion of arable land to pasture would lead to a reduction in soil loss across Europe. However, Panagos et al. (2015) also highlighted that the susceptibility of pastures to erosion varies by region, depending on the climatic conditions and the characteristics of the pasture. Grasslands where grazing is significant can have an increase in sediment export and turbidity (Trimble and Mendel 1995, Zaimes et al. 2004). If cattle grazing in the grassland areas of the studied catchments is significant, such areas would contribute to higher values of turbidity.

Mean 72-h rainfall prior to sampling was the only explanatory site-scale variable that was significant in the global model and all the city-scale models. Volume and intensity of rainfall events are critical factors for soil erosion and sediment transport (Ahearn et al. 2005, Lawler et al. 2006b, Li et al. 2009, Fortin et al. 2015), therefore, an increase in turbidity with elevated precipitations was expected. However, rainfall and water quality in urbanized catchments have a complex relationship (McGrane et al. 2017). In our study, the relationship between rainfall and turbidity was inconsistent across models. In spite of the heterogeneous response of turbidity to rainfall, turbidity plateaued at high rainfall

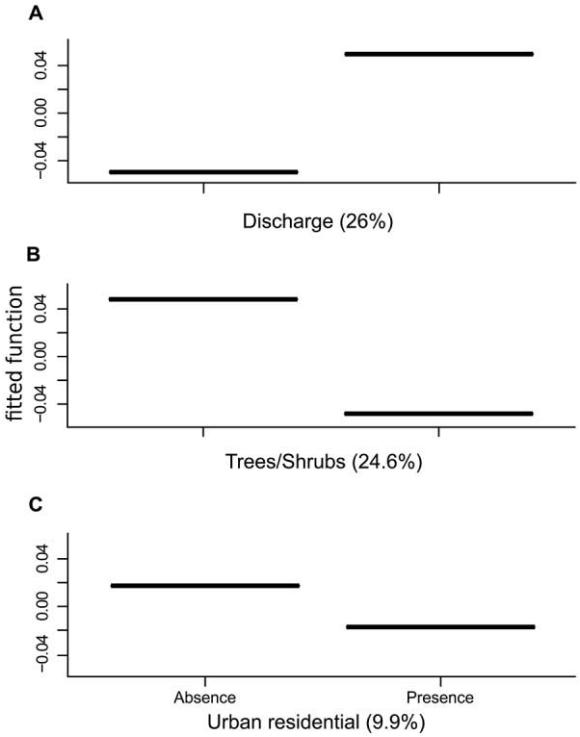


Figure 11. Partial plots for the boosted regression trees (BRT) model for data collected by citizen scientists only. Plots are ordered by importance of the covariates in the final boosted regression tree model (Table 3). Discharge (point source discharge) (A), trees/shrubs (riverbank trees and shrubs) (B), urban residential (C).

values in all models, suggesting dilution effects. This effect is often seen in combined sewer systems (systems that route both sewage and storm water together), where an initial increase in particulate concentrations from runoff is followed by a decrease related to dilution (Gupta and Saul 1996, Morgan et al. 2017). However, this 'first flush' action does not always occur in urbanized rivers (Deletic and Maksimovic 1989, Lawler et al. 2006a). Another explanation for the heterogeneity in response to rainfall could be that relatively few high-precipitation events were included in the dataset. A limitation of citizen scientist acquired data is that sampling in the rain, or soon after high-rain events, is often avoided by choice or design (Thornhill et al. 2016). Approximately 80% of the samples in this study were collected in periods with no or little precipitation (<0.25 mm/h), and only 1% of the samples were collected on high precipitation days (>0.83 mm/h).

Cropland land use did not significantly explain turbidity at global scale but was an important factor in 4 of the 6 city models. Agricultural activity is typically associated with high erosion rates (Allan et al. 1997, Harding et al. 1999, García-Ruiz et al. 2015), but we only found a significant increase of turbidity with increased cropland in Vancouver. On the other hand, São Paulo, Guangzhou, and Buenos Aires showed the opposite trend with lower values of turbidity when the proportion of coverage of croplands in the catchments was higher. These contradictory results could be related to the variability in the erosion rates from agricultural areas observed by García-Ruiz et al. (2015) and the variety of impacts that agriculture can have on water bodies (Harding et al. 1999). Also, the percentage of cropland in the catchment does not consider the proximity of agricultural activities to receiving rivers or consider the presence of any interceding buffer areas. Both factors can influence sediment load (Omernik et al. 1981, Buck et al. 2004).

## Contribution of citizen scientists to the explanation of turbidity patterns

Our study showed that site-scale factors had less power than catchment-scale factors in predicting turbidity levels. Several other studies have also found that landscape characteristics had greater influence on water quality than the river buffer areas (Hunsaker and Levine 1995, Sliva and Williams 2001). However, other studies have found that TSS and turbidity were better explained by land-cover characteristics at the riparian-buffer scale (Johnson et al. 1997, Chang 2008). Nevertheless, the information collected by citizen scientists at the site scale explained 11.6% of the variability in turbidity. This site-scale information points to the possibility that local actions could be implemented to reduce turbidity.

The presence of locally observed point source discharges was associated with higher turbidity. These point source discharges in urban areas had different origins, such as industrial, residential, roads, septic systems, or storm water drainage pipes. The presence of storm water drainage inflows, the most common type of discharge present in most urban areas, has been shown to increase turbidity during precipitation events (Lawler et al. 2006b, Wenger et al. 2009). This finding supports earlier studies that showed point-source discharges identified by citizen scientists can be a useful predictor of water quality (e.g., nutrients, Loiselle et al. 2016).

It is well established that the presence of riverbank vegetation can reduce particulate concentrations through multiple mechanisms, such as decreasing the severity of streambed and riverbank erosion (Hubble et al. 2010, Gumiere et al. 2011) and improving the retention of suspended particles in the catchment by reducing in-flow velocity (Johnson et al. 1997, Sliva and Williams 2001). Riverbank vegetation around urban streams is usually structurally limited and managed for amenity value more than for ecological value (Wenger et al. 2009). Further, several studies have suggested that even intact urban riparian zones have a reduced functionality, as most stormwater runoff is routed directly into streams via the stormwater network (Roy et al. 2005, 2006). Still, the present study indicates that riverbank trees and shrubs reduced turbidity across a range of climate and urban settings, supporting the results of previous studies (Parkyn et al. 2003, Wenger et al. 2009).

When urban residential cover was the main land use around a sampling site, turbidity was lower than in sampling sites where urban residential cover was not the main land use. In urban residential areas, bare land is often covered with impervious surfaces (Moreno Madriñán et al. 2012), so less soil may be available to erode and contribute to turbidity. Additionally, the density of the storm water network is often higher in urban residential areas, which disconnects the river channel from nearby overland runoff. These factors may explain why we found a stronger relationship between turbidity and point source discharges than with urban residential land cover.

## Conclusions

The growing availability of both catchment-scale data and citizen scientist generated site-scale data allows for new insights into the turbidity dynamics at both global and local scales. This information is fundamental to improving management and decision making in areas where the urbanization of peri-urban and natural areas is significant. This study provided evidence that a multiscale approach can be used, addressing site-scale factors through local actions, in combination with catchment planning to reduce turbidity levels and maintain a functioning river environment. For example, creating and maintaining vegetated riverbanks in cities and promoting better management of point-source discharges could contribute to the reduction of turbidity. With further refinement, community-based monitoring (e.g., citizen science) is a promising method to gather site-scale information in large urban areas and help evaluate and improve river

conditions as well as improve the knowledge of local communities.

Furthermore, the participation of local communities in river management is important for creating more integrated catchment management plans (Downs et al. 1991, Parkyn et al. 2003). Indeed, a previous study observed that "The success of any attempt to improve the ecological condition of streams in urban areas will largely depend on human attitudes and behavior" (Booth 2005). Therefore, integrating social and ecological aspects of catchment management, through citizen science, has both direct management benefits from increased site-scale information, as well as indirect management benefits from an improved awareness of the local conditions.

## ACKNOWLEDGEMENTS

Author contributions: LMC contributed to the study design, carried out the statistical analyses, led authorship of the manuscript, and produced figures and tables. IT contributed to the study design, helped with data analysis, discussed the results, and made comments and edits on the manuscript. SL contributed to the study design, discussed the results, and made comments on the manuscript. EH helped to frame and edit the manuscript.

We thank HSBC Bank for the financial support of the Fresh-Water Watch, under the scope of the HSBC Water Program. We sincerely acknowledge the efforts of the citizen scientists who were active in the program and the participating project scientists in each study city. Special thanks are extended to editor Charles P. Hawkins, technical editor Katherine M. Sirianni, and 2 anonymous referees for their feedback and helpful comments on the manuscript. The NCEP Reanalysis 2 data were obtained from the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (www.esrl.noaa.gov/psd/).

## LITERATURE CITED

Ahearn, D. S., R. W. Sheibley, R. A. Dahlgren, M. Anderson, J. Johnson, and K. W. Tate. 2005. Land use and land cover influence on water quality in the last free-flowing river draining the western Sierra Nevada, California. Journal of Hydrology 313:234–247.

Allan, J. D., D. L. Erickson, and J. Fay. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37:149–161.

Bernhardt, E. S., and M. A. Palmer. 2007. Restoring streams in an urbanizing world. Freshwater Biology 52:738–751.

Booth, D. B. 2005. Challenges and prospects for restoring urban streams: a perspective from the Pacific Northwest of North America. Journal of the North American Benthological Society 24:724–737.

Booth, D. B., and C. R. Jackson. 1997. Urbanization of aquatic systems: degradation thresholds, stormwater detection, and the limits of mitigation. Journal of the American Water Resources Association 33:1077–1090.

Booth, D. B., A. H. Roy, B. Smith, and K. Capps. 2016. Global perspectives on the urban stream syndrome. Journal of Freshwater Science 36:1–11.

Brett, M. T., G. B. Arhonditsis, S. E. Mueller, D. M. Hartley, J. D. Frodge, and D. E. Funke. 2005. Non-point-source impacts on stream nutrient concentrations along a forest to urban gradient. Environmental Management 35:330–342.

Bruhn, L., and P. Soranno. 2005. Long term (1974–2001) volunteer monitoring of water clarity trends in Michigan Lakes and their relation to ecoregion and land use/cover. Lake and Reservoir Management 21:10–23.

Buck, O., D. K. Niyogi, and C. R. Townsend. 2004. Scale-dependence of land use effects on water quality of streams in agricultural catchments. Environmental Pollution 130:287–299.

Canfield, D. E., C. D. Brown, R. W. Bachmann, and M. V. Hoyer. 2002. Volunteer lake monitoring: testing the reliability of data collected by the Florida LAKEWATCH Program. Lake and Reservoir Management 18:1–9.

Capps, K. A., C. N. Bentsen, and A. Ramírez. 2016. Poverty, urbanization, and environmental degradation: urban streams in the developing world. Freshwater Science 35:429–435.

Castilla, E. P., D. G. F. Cunha, F. W. F. Lee, S. Loiselle, K. C. Ho, and C. Hall. 2015. Quantification of phytoplankton bloom dynamics by citizen scientists in urban and peri-urban environments. Environmental Monitoring and Assessment 187:690.

Chang, H. 2008. Spatial analysis of water quality trends in the Han River basin, South Korea. Water Research 42:3285–3304.

CIESIN (Center for International Earth Science Information Network). 2016. Documentation for the Gridded Population of the World, version 4 (GPWv4). Palisades NY: NASA Socioeconomic Data and Applications Center (SEDAC). Columbia University Press, New York.

Cohen, B. 2006. Urbanization in developing countries: current trends, future projections, and key challenges for sustainability. Technology in Society 28:63–80.

Cunha, D. G. F., S. P. Casali, P. B. de Falco, I. Thornhill, and S. A. Loiselle. 2017. The contribution of volunteer-based monitoring data to the assessment of harmful phytoplankton blooms in Brazilian urban streams. Science of the Total Environment 584–585:586–594.

Davies, N. M., R. H. Norris, and M. C. Thoms. 2000. Prediction and assessment of local stream habitat features using large-scale catchment characteristics. Freshwater Biology 45:343–369.

Deletic, A., and C. Maksimovic. 1989. Evaluation of water quality factors in storm runoff from paved areas. Journal of environmental engineering 124:869–879.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, and S. Lautenbach. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46.

Downs, P. W., K. J. Gregory, and A. Brookes. 1991. How integrated is river basin management? Environmental Management 15:299–309.

Dunne, T., and L. B. Leopold. 1978. Water in environmental planning. W. H. Freeman, San Francisco, California.

Eigenbrod, F., V. A. Bell, H. N. Davies, A. Heinemeyer, P. R. Armsworth, and K. J. Gaston. 2011. The impact of projected increases in urbanization on ecosystem services. Proceedings of the Royal Society of London Series B: Biological Sciences 278:3201–3208.

Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77:802–813.

Fortin, G., M. LeBlanc, S. Schiavone, O. Chouinard, and A. Utzschneider. 2015. Local perceptions, RUSLEFAC mapping, and field

results: the sediment budget of Cocagne River, New Brunswick, Canada. Environmental Management 55:113–127.

Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. The Annals of Statistics 29:1189–232.

Friedman, J. H. 2002. Stochastic gradient boosting. Computational Statistics and Data Analysis 38:367–378.

García-Ruiz, J. M., S. Beguería, E. Nadal-Romero, J. C. González-Hidalgo, N. Lana-Renault, and Y. Sanjuán. 2015. A meta-analysis of soil erosion rates across the world. Geomorphology 239:160–173.

Gumiere, S. J., Y. Le Bissonnais, D. Raclot, and B. Cheviron. 2011. Vegetated filter effects on sedimentological connectivity of agricultural catchments in erosion modelling: a review. Earth Surface Processes and Landforms 36:3–19.

Gupta, K., and A. Saul. 1996. Suspended solids in combined sewer flows. Water Science and Technology 33:93–99.

Harding, J. S., R. G. Young, J. W. Hayes, K. A. Shearer, and J. D. Stark. 1999. Changes in agricultural intensity and river health along a river continuum. Freshwater Biology 42:345–357.

Harrel, F. E. Jr., and C. Dupont, 2015. Hmisc: Harrell miscellaneous. R package version 3.15-0.

Henley, W. F., M. A. Patterson, R. J. Neves, and A. D. Lemly. 2000. Effects of sedimentation and turbidity on lotic food webs: a concise review for natural resource managers. Reviews in Fisheries Science 8:125–139.

Hubble, T. C. T., B. B. Docker, and I. D. Rutherfurd. 2010. The role of riparian trees in maintaining riverbank stability: A review of Australian experience and practice. Ecological Engineering 36:292–304.

Huey, G. M., and M. L. Meyer. 2010. Turbidity as an indicator of water quality in diverse watersheds of the Upper Pecos River Basin. Water 2:273–284.

Hunsaker, C., and D. Levine. 1995. Hierarchical Approaches to the study of water quality in rivers. BioScience 45:193–203.

Johnson, L., C. Richards, G. Host, and J. Arthur. 1997. Landscape influences on water chemistry in Midwestern stream ecosystems. Freshwater Biology 37:193–208.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter. 2002. NCEP–DOE AMIP-II Reanalysis (R-2). Bulletin of the American Meteorological Society 83:1631–1643.

Kemp, M. U., E. E. Van Loon, J. Shamoun-Baranes, and W. Bouten. 2012. RNCEP: global weather and climate data at your fingertips 3:65–70.

Latham, J., R. Cumani, I. Rosati, and M. Bloise. 2014. FAO Global Land Cover (GLC-SHARE) Beta-Release 1.0 Database, Land and Water Division. (Available from: http://www.fao.org/geonetwork/srv/en/main.home?uuid=ba4526fd-cdbf-4028-a1bd-5a559c4bff38)

Lathrop, R. C., S. R. Carpenter, and L. G. Rudstam. 1996. Water clarity in Lake Mendota since 1900: responses to differing levels of nutrients and herbivory. Canadian Journal of Fisheries and Aquatic Sciences 53:2250–2261.

Lawler, D. M., I. D. L. Foster, G. E. Petts, S. Harper, and I. P. Morrissey. 2006a. Suspended sediment dynamics for June storm events in the urbanized River Tame, UK. Sediment Dynamics and the Hydromorphology of Fluvial Systems 306:96–103.

Lawler, D. M., G. E. Petts, I. D. L. Foster, and S. Harper. 2006b. Turbidity dynamics during spring storm events in an urban headwater river system: the Upper Tame, West Midlands, UK. Science of the Total Environment 360:109–126.

Lehner, B., K. Verdin, and A. Jarvis. 2008. New global hydrography derived from spaceborne elevation data. Eos, Transactions American Geophysical Union 89:93–94.

Li, S., S. Gu, W. Liu, H. Han, and Q. Zhang. 2008. Water quality in relation to land use and land cover in the upper Han River Basin, China. Catena 75:216–222.

Li, S., S. Gu, X. Tan, and Q. Zhang. 2009. Water quality in the upper Han River basin, China: the impacts of land use/land cover in riparian buffer zone. Journal of Hazardous Materials 165: 317–324.

Loiselle, S. A., D. G. F. Cunha, S. Shupe, E. Valiente, L. Rocha, E. Heasley, P. Pérez Belmont, and A. Baruch. 2016. Micro and macroscale drivers of nutrient concentrations in urban streams in South, Central and North America. PLoS ONE 11(9):e0162684.

Loiselle, S. A., P. C. Frost, E. Turak, and I. Thornhill. 2017. Citizen scientists supporting environmental research priorities. Science of the Total Environment 598:937.

Lottig, N. R., T. Wagner, E. Norton Henry, K. Spence Cheruvelil, K. E. Webster, J. A. Downing, and C. A. Stow. 2014. Long-term citizen-collected data reveal geographical patterns and temporal trends in lake water clarity. PLoS ONE 9:e95769.

McGrane, S. J., M. G. Hutchins, J. D. Miller, G. Bussi, T. R. Kjeldsen, and M. Loewenthal. 2017. During a winter of storms in a small UK catchment, hydrology and water quality responses follow a clear rural-urban gradient. Journal of Hydrology 545:463–477.

Meyer, J. L., M. J. Paul, and W. K. Taulbee. 2005. Stream ecosystem function in urbanizing landscapes. Journal of the North American Benthological Society 24:602–612.

Moreno Madriñán, M. J., M. Z. Al-Hamdan, D. L. Rickman, and J. Ye. 2012. Relationship between watershed land-cover/land-use change and water turbidity status of Tampa Bay major tributaries, Florida, USA. Water, Air, and Soil Pollution 223:2093–2109.

Morgan, D., P. Johnston, K. Osei, and L. Gill. 2017. The influence of particle size on the first flush strength of urban stormwater runoff. Water Science and Technology 76:2140–2149.

Müller, D., P. J. Leitão, and T. Sikor. 2013. Comparing the determinants of cropland abandonment in Albania and Romania using boosted regression trees. Agricultural Systems 117:66–77.

Myre, E., and R. Shaw. 2006. The turbidity tube: simple and accurate measurement of turbidity in the field. Michigan Technological University, Houghton, Michigan.

Obrecht, D. V., M. Milanick, B. D. Perkins, D. Ready, and J. R. Jones. 1998. Evaluation of Data Generated from Lake Samples Collected by Volunteers. Lake and Reservoir Management 14: 21–27.

O'Driscoll, M., S. Clinton, A. Jefferson, A. Manda, and S. McMillan. 2010. Urbanization Effects on Watershed Hydrology and In-Stream Processes in the Southern United States. Water 2:605–648.

Olson, D. M., E. Dinerstein, E. D. Wikramanayake, N. D. Burgess, G. V. N. Powell, E. C. Underwood, J. A. D'Amico, I. Itoua, H. E. Strand, J. C. Morrison, C. J. Loucks, T. F. Allnutt, T. H. Ricketts, Y. Kura, J. F. Lamoureux, W. W. Wettengel, P. Hedao, and K. R. Kassem. 2001. Terrestrial ecoregions of the world: a new map of life on earth: a new global map of terrestrial ecoregions pro-

vides an innovative tool for conserving biodiversity. BioScience 51:933–938.

Olden, J. D., J. J. Lawler, and N. L. Poff. 2008. Machine learning methods without tears: a primer for ecologists. The Quarterly Review of Biology 83:171–193.

Omernik, J. M., A. R. Abernathy, and L. M. Male. 1981. Stream nutrient levels and proximity of agricultural and forest land to streams: some relationships. Journal of Soil and Water Conservation 36:227–231.

Panagos, P., P. Borrelli, K. Meusburger, C. Alewell, E. Lugato, and L. Montanarella. 2015. Estimating the soil erosion cover-management factor at the European scale. Land Use Policy 48:38–50.

Parkyn, S. M., R. J. Davies-Colley, N. J. Halliday, K. J. Costley, and G. F. Croker. 2003. Planted riparian buffer zones in New Zealand: do they live up to expectations? Restoration Ecology 11:436–447.

Paul, M. J., and J. L. Meyer. 2001. Streams in the urban landscape. Annual Review of Ecology and Systematics 32:333–365.

Paule-Mercado, M. A., J. S. Ventura, S. A. Memon, D. Jahng, J. H. Kang, and C. H. Lee. 2016. Monitoring and predicting the fecal indicator bacteria concentrations from agricultural, mixed land use and urban stormwater runoff. Science of the Total Environment 550:1171–1181.

Pearson, R. 2017. Interpreting Predictive Models Using Partial Dependence Plots. https://cran.r-project.org/web/packages/datarobot/vignettes/PartialDependence.html.

Peters, N. E. 2009. Effects of urbanization on stream water quality in the city of Atlanta, Georgia, USA. Hydrological Processes 23:2860–2878.

Preisendorfer, R. W. 1986. Secchi disk science: Visual optics of natural waters 1. Limnology and Oceanography 31:909–926.

R Core Development Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rains, M. C., R. A. Dahlgren, G. E. Fogg, T. Harter, and R. J. Williamson. 2008. Geological control of physical and chemical hydrology in California vernal pools. Wetlands 28:347–362.

Ridgeway, G. 2015. Generalized boosted regression models. R package version 2.1.1.

Roy, A. H., M. C. Freeman, B. J. Freeman, S. J. Wenger, W. E. Ensign, and J. L. Meyer. 2005. Investigating hydrologic alteration as a mechanism of fish assemblage shifts in urbanizing streams. Journal of the North American Benthological Society 24:656–678.

Roy, A. H., M. C. Freeman, B. J. Freeman, S. J. Wenger, J. L. Meyer, and W. E. Ensign. 2006. Importance of riparian forests in urban catchments contingent on sediment and hydrologic regimes. Environmental Management 37:523–539.

Roy, A. H., A. D. Rosemond, M. J. Paul, D. S. Leigh, and J. B. Wallace. 2003. Stream macroinvertebrate response to catchment urbanization (Georgia, U.S.A.). Freshwater Biology 48:329–346.

Rügner, H., M. Schwientek, B. Beckingham, B. Kuch, and P. Grathwohl. 2013. Turbidity as a proxy for total suspended solids (TSS) and particle facilitated pollutant transport in catchments. Environmental Earth Sciences 69:373–380.

Ryan, P. A. 1991. Environmental effects of sediment on New Zealand streams: a review. New Zealand Journal of Marine and Freshwater Research 25:207–221.

Sliva, L., and D. D. Williams. 2001. Buffer zone versus whole catchment approaches to studying land use impact on river water quality. Water Research 35:3462–72.

Thornhill, I., S. Loiselle, K. Lind, and D. Ophof. 2016. The citizen science opportunity for researchers and agencies. BioScience 66:720–721.

Trimble, S. W., and A. C. Mendel. 1995. The cow as a geomorphic agent—a critical review. Geomorphology 13:233–253.

Tyler, J. E. 1968. The Secchi disk. Limnology and Oceanography 13:1–6.

UNPD (United Nations Procurement Division). 2014. World urbanization prospects. The 2014 revision.

Vairavamoorthy, K., S. D. Gorantiwar, and A. Pathirana. 2008. Managing urban water supplies in developing countries—climate change and water scarcity scenarios. Physics and Chemistry of the Earth 33:330–339.

Walsh, C. J., T. D. Fletcher, and A. R. Ladson. 2005a. Stream restoration in urban catchments through redesigning stormwater systems: looking to the catchment to save the stream. Journal of the North American Benthological Society 24:690–705.

Walsh, C. J., A. H. Roy, J. W. Feminella, P. D. Cottingham, P. M. Groffman, and R. P. Morgan II. 2005b. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24:706–723.

Walters, D. M., D. S. Leigh, and A. B. Bearden. 2003. Urbanization, sedimentation, and the homogenization of fish assemblages in the Etowah River Basin, USA. Hydrobiologia 494:5–10.

Walters, D. M., A. H. Roy, and D. S. Leigh. 2009. Environmental indicators of macroinvertebrate and fish assemblage integrity in urbanizing watersheds. Ecological Indicators 9:1222–1233.

Wenger, S. J., A. H. Roy, C. R. Jackson, E. S. Bernhardt, T. L. Carter, S. Filoso, C. A. Gibson, W. C. Hession, S. S. Kaushal, E. Martí, J. L. Meyer, M. A. Palmer, M. J. Paul, A. H. Purcell, A. Ramírez, A. D. Rosemond, K. A. Schofield, E. B. Sudduth, and C. J. Walsh. 2009. Twenty-six key research questions in urban stream ecology: an assessment of the state of the science. Journal of the North American Benthological Society 28:1080–1098.

Wernand, M. R. 2010. On the history of the Secchi disc. Journal of the European Optical Society: Rapid Publications 5:10013s.

West, A. O., and J. T. Scott. 2016. Black disk visibility, turbidity, and total suspended solids in rivers: a comparative evaluation. Limnology and Oceanography: Methods 14:658–667.

Zaimes, G. N., R. C. Schultz, and T. M. Isenhart. 2004. Stream bank erosion adjacent to riparian forest buffers, row crop fields, and continuously-grazed pastures along Bear Creek in central Iowa. Journal of Soil and Water Conservation 50:19–27.

Zhang, Q., Z. Li, G. Zeng, J. Li, Y. Fang, Q. Yuan, Y. Wang, and F. Ye. 2009. Assessment of surface water quality using multivariate statistical techniques in red soil hilly region: a case study of Xiangjiang watershed, China. Environmental Monitoring and Assessment 152:123–131.

Zhang, Y., R. Ma, M. Hu, J. Luo, J. Li, and Q. Liang. 2017. Combining citizen science and land use data to identify drivers of eutrophication in the Huangpu River system. Science of the Total Environment 584–585:651–664.

Zhou, T., J. Wu, and S. Peng. 2012. Assessing the effects of landscape pattern on river water quality at multiple scales: a case study of the Dongjiang River watershed, China. Ecological Indicators 23:166–175.